HOW TO CASTf90

Sabine Radanovics

November 16, 2015

1 What does it do?

CASTf90 first downloads fields from NCEP reanalysis (sea level pressure, slp, as default) and then searches for a given simulation period the most similar cases within a given data base period according to a given distance measure. Finally it writes the N most similar days including the calculated distances for them to an output file.

2 Requirements

CASTf90 is developed for Linux systems. It might run on macOSX, but this was never tested.

CASTf90 requires:

- cdo (climate data operators)
- nco (netCDF operators)
- netCDF with Fortran90 interface
- a Fortran95 compiler
- lapack95 libraries

3 Get the code

The code is located in a cvs repository situated on the on the common file system (make sure to be logged in on obelix...), so the first step is to checkout a so called working copy from the repository. To this end:

1. create a directory where you want the code to be located

   ```
   mkdir Mydir
   ```

   and change to this directory

   ```
   cd Mydir
   ```

2. set the environment variable CVSROOT to /home/users/sradanov/cvs. The command depends on the active shell but might be something like

```
export CVSROOT=/home/users/sradanov/cvs
```

Now cvs knows where to look for the repository.

3. checkout the working copy from the repository with

```
cvs checkout Analogue/RSdev
```

Now you should have some fortran90 source files, some shell scripts, an example pbsscript and a makefile in your directory. (See section 3.1 for a complete list.)

From time to time you may run

```
cvs status
```

to see if the source files are still all Up-to-date. If this is not the case, you may run

```
cvs update
```

to get the latest version of the source files and scripts. Attention! If .f90 sources were updated the program has to be recompiled! (see section 4 for details). If run_analogs_case.sh has been updated, don't forget to change the variable *sourcedir*.

For further details on cvs and its commands please refer to the cvs manual.

3.1 List of scripts and source files

After checkout the following files should be present:

Shellscripts

- run_analogs_case.sh
- getNCEP_slp.sh
- retrieve.sh

Source files

- analogue.f90
- config.f90
- distance.f90
- eofs.f90
- read.f90
- routines.f90

Makefile

- Makefile

Example pbsscript

- pbsscript

4 Compile the program

CASTf90 comes with a makefile to compile the fortran90 sources, however, it is assumed that netcdf libraries with fortran90 interface and the lapack library with fortran95 interface are installed. If you are working on "obelix" you just have to load the netcdf module using

```
module load netcdf/4p
```

before running

```
make -f Makefile
```

Elsewhere please make sure that netcdf **AND** its fortran90 interface are installed and that INCLUDEPATH in the makefile is the directory containing netcdf.mod, and LIBPATH the one containing libnetcdff.

The intel compiler already comes with the -mkl option that gives access to the lapack library, however, the lapack95 interface may need to be installed separately. INCLUDEPATH2 and LIBPATH2 contain the .mod and the lib files for lapack95 respectively.

For other compilers the entire lapack library may need to be compiled and compiler flags will need to be adapted. (The most important is probably the one enabeling OpenMP.)

5 Run the script

Before running the script the first time or after an update, please verify that the variable *sourcedir* in run_analogs_case.sh (line 37) is set to the directory containing the source code.

To anable parallel calculation the OMP_NUM_THREADS environement variable has to be set. The command depends on the active shell, but might be something like

```
export OMP_NUM_THREADS=12
```

The number has to be adjusted according to the available ressources. When running through a batch system, this should be done in the batch script (see section 5.1 for details).

Then to start the procedure, run

```
./run_analogs_case.sh
```

run_analogs_case.sh accepts a number of options:

- Options to set I/O paths:
  - **-P**<*path to base data*>
  - **-p**<*path to simulation data*>
  - **-o**<*path to output file*>

  Paths have to end with a slash. Example:

  ```
  ./run_analogs_case.sh -P/home/scratch01/sradanov/A2C2/NCEP/
  ```

- Options related to the geographical region of interest:
  - **-R**<*region name*> Known region names are NA (North Atlantic) and NHmid (Northern Hemisphere midlatitudes). This region is used at the time of data download. Using a limited number of large regions for downloads avoids to store a version of the data for every domain one might want to use and to skip the download step if the data is already present. The -D option allows then to specify the spatial domain the analogues are actually calculated upon. New regions can be added in the script retrieve.sh from line 20 on.
  - **-D**<*lonmin*>,<*lonmax*>,<*latmin*>,<*latmax*> predictor domain, that is the spatial domain for which the analogues should be calculated. The domain has to be inside the region defined in option -R. Example:

    ```
    ./run_analogs_case.sh -RNA -D-20.0,50.0,22.5,70.0
    ```

- Time period selection:
  - **-S**<*YYYY-MM-DD*>,<*YYYY-MM-DD*> Simulation period
  - **-B**<*YYYY-MM-DD*>,<*YYYY-MM-DD*> Database/archive period (period from which the analogues are chosen)

  Example:

  ```
  ./run_analogs_case.sh -S2013-12-01,2014-02-28 -B1950-01-01,1979-12-31
  ```

- Anomalie options
  - **-N**<*mode*> while <mode> is one of the following:
    - *none* analogues are calculated for raw fields.
    - *base* anomalies are calculated with respect to a smoothed seasonal cycle calculated from the database/archive period. The seasonal cycle is first calculated as day of year average over the years in the database period and then smoothed using a weighted moving average. For the smoothing window parameter see option -m.
    - *sim* anomalies are calculated with respect to a smoothed seasonal cycle calculated over the simulation period. (caution: make sure that the simulation is sufficiently long for meaningful cycle calculation)
    - *own* anomalies are calculated for each dataset (database and simulation) with respect to its own smoothed seasonal cycle. This option might be useful when simulation and database sets are from different sources and one want to get rid of the mean difference between the two data sets. (caution: change signals that manifest themselfs in the mean difference will probably be lost)

- **-m**<*numberofdays*> number of days (preferably an odd number) to calculate weighted moving average for seasonal cycle smoothing. Weights are linearly decreasing towards both ends of the interval.

Example:

```
./run_analogs_case.sh -Nbase -m91
```

- Analogue calculation options:
  - **-v**<*varname*> Name of the NCEP field to download. The name has to be the same as in the filename in the NCEP database. Example for precipitable water:

    ```
    ./run_analogs_case.sh  -vpr_wtr.eatm
    ```

  - **-w**<*numberofdays*> Number of days of the year ± around the target day to consider as candidates. This parameter allows a seasonal restriction of the analogue search. For no seasonal restriction choose 183.
  - **-d**<*distance*> Name of the distance to use for analogue calculation. Supported distances are (in increasing order of complexity):
    - *euclidean* (or equivalently *rms*, *rmse*)
    - *mahalanobis*
    - *of* (or equivalently *opticalflow*) Which is a displacement and amplitude difference similar to Keil and Craig (2009) derived using the optical flow field deformation technique (Marzban and Sandgathe, 2010). (Experimental...)

    Attention: the more complex the distance the longer the computation time. The computation time strongly depends on the domain size and the archive length!
  - **-C**<*numberofdays*> Distances can be averaged over a *number of consecutive days* in order to find situations with similar time evolution.
  - **-n**<*numberofanalogs*> Number of closest analogue dates to write to output.
  - **-c**<*logical*> TRUE if rank correlation should be calculated as an additional diagnostic (written to the output file), FALSE if not.

Example:

```
./run_analogs_case.sh -w30 -dmahalanobis -C3 -n25 -cTRUE
```

- **-help** lists the options and default settings in the terminal and exits.

5.1 Batch system

If the run_analogs_case.sh script is submitted to a batch system (qsub command), the options are interpreted as options for the qsub command rather than as options for the run_analogs_case.sh script. A workaround is to run run_analogs_case.sh inside an other script which is then submitted to the batch system. *pbsscript* is an example for such a script.

6 Output

The output is a multi-column text file containing:

- the date of the analyzed day (in the form yyyymmdd)
- the dates of the n best analogues (in the form yyyymmdd)
- the value of the minimised distance (to determine the analogues).
- the value of the spatial rank correlations between the n analogues and the analysed field if the -c option is TRUE.

The output file can serve as input for the AnaWEGE weather generator (Yiou, 2014).

7 Legal Issues (Disclaimer)

This software program is protected by the intellectual property right. Its owners grant to you, free of charge, the right to use this software program exclusively for research purposes provided you clearly specify the name of the software program and the here above copyright notice in any publication, written by you, related to research results obtained and/or consolidated, in whole or in part, from the use of the software program. If source code of this software program has been delivered, you have to inform us of any modification of the software program by sending an email to pascal.yiou@lsce.ipsl.fr. Any distribution of this software program is forbidden.

For any other use, notably for any commercial use, you have to subscribe a specific license. In this case, send an email to pascal.yiou@lsce.ipsl.fr .

7.1 Liability

The owners liability shall not be incurred as a result of in particular (i) loss due the Licensee's total or partial failure to fulfill its obligations, (ii) direct or consequential loss that is suffered by the Licensee due to the use or performance of the software, and (iii) more generally, any consequential loss.

7.2 Warranty

You acknowledge that the scientific and technical state-of-the-art when the software program was distributed did not enable all possible uses to be tested and verified, nor for the presence of possible defects to be detected. In this respect, the your attention has been drawn to the risks associated with loading, using,

modifying and/or developing and reproducing the software which are reserved for experienced users. You shall be responsible for verifying, by any or all means, the suitability of the product for your requirements, its good working order, and for ensuring that it shall not cause damage to either persons or properties. The owners hereby represents, in good faith, that they are entitled to grant all the rights over the software program. You acknowledge that the software program is supplied "as is" by the owners without any other express or tacit warranty, other than that provided for here above and, in particular, without any warranty as to its commercial value, its secured, safe, innovative or relevant nature. Specifically, the owners do not warrant that the software program is free from any error, that it will operate without interruption, that it will be compatible with your own equipment and software configuration, nor that it will meet your requirements. The owners do not either expressly or tacitly warrant that the software program does not infringe any third party intellectual property right relating to a patent, software or any other property right. Therefore, the owners disclaim any and all liability towards you arising out of any or all proceedings for infringement that may be instituted in respect of the use, modification and redistribution of the software program. Nevertheless, should such proceedings be instituted against you, the owners shall provide it with technical and legal assistance for your defense. The owners disclaim any and all liability as regards your use of the name of the software program. No warranty is given as regards the existence of prior rights over the name of the software program or as regards the existence of a trademark.

References

Keil, C. and Craig, G. C. (2009). A displacement and amplitude score employing an optical flow technique. *Weather and Forecasting*, 24:1297–1308.

Marzban, C. and Sandgathe, S. (2010). Optical flow for verification. *Weather and Forecasting*, 25:1479–1494.

Yiou, P. (2014). Anawege: a weather generator based on analogues of atmospheric circulation. *Geoscientific Model Development*, 7(2):531–543.